Automated calibration, meta-data, and trust

David W. Hogg
Center for Cosmology and Particle Physics, New York University

2009 February 6

non-text search

- all (important) image search is based on associated text
- want to search the content of the *image pixels* themselves
  - holy grail of computer vision
  - enormous commercial value
  - obviates *trusting* the associated meta-data
- we have solved this in one *tiny* domain

Gmail - Infor… | General info… | Setting up a … | New York Un… | Google Cale… | How to Take… | Astrometry.…

Astrometry.net: Web Edition

Required Information

Identification

We need your email address to send you your results.

Email: [ ]

Name: [ ]

☐ Remember me

Field to solve: select one of the following:

In order to solve your field, we need either an image of your field (as a FITS, JPEG, GIF, PNG file or a URL); OR a list of star positions (as a FITS binary table or text file).

○ Image file: [ ] Browse...

Your JPEG, GIF, PNG, or FITS image. Must be less than 250 MB!

○ Image URL: http://

The URL of a JPEG, GIF, PNG, or FITS image. Must be less than 250 MB!

○ FITS binary table:

Must contain a BINTABLE of detected objects, with X and Y pixel positions in "D" (double) or "E" (float) columns, with one object per row. The rows must be sorted by brightness with the brightest object first.

- FITS file: [ ] Browse...
- X Column name: X
- Y Column name: Y

○ Text file: [ ] Browse...

A text file, containing two columns of digits separated

Optional Settings

Parity of your image

Flipping an image reverses its "parity". If you point a digital camera at the sky and submit the JPEG, it probably has negative parity. If you have a FITS image, it probably has positive parity. Selecting the right parity will make the solving process run faster, but if in doubt just try both.

◉ Try both parities simultaneously

○ Positive parity image

○ Negative parity image

Index to use

We use indexes to solve your fields; an index contains "landmarks" for the sky. The solving process works best when the "landmarks" are a little smaller than the size of the field to be solved. We can automatically choose a suitable index based on your pixel scale estimate, or you can select one here.

Automatic (based on image scale)

Star positional error

When we find a matching "landmark", we check to see how many of the stars in your field match up with stars we know about. To do this, we need to know how much a star in your field could have moved from where it should be.

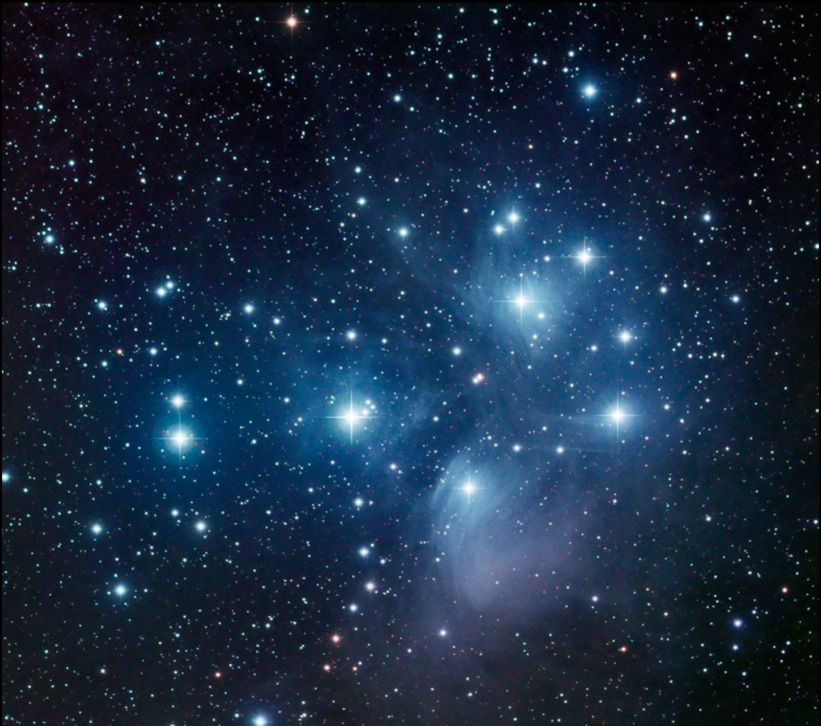
Positional error, in pixels: 1

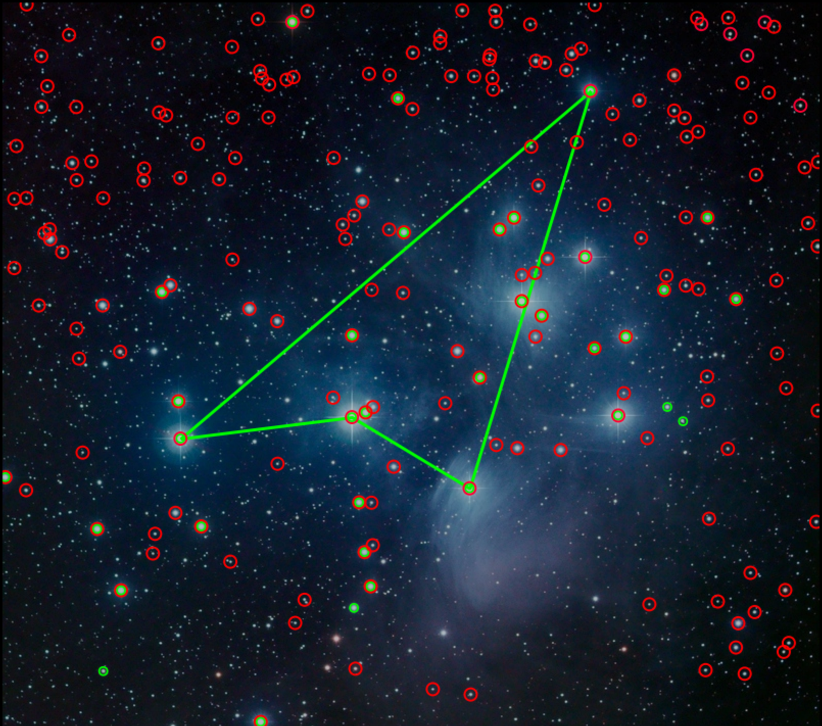
Tweak

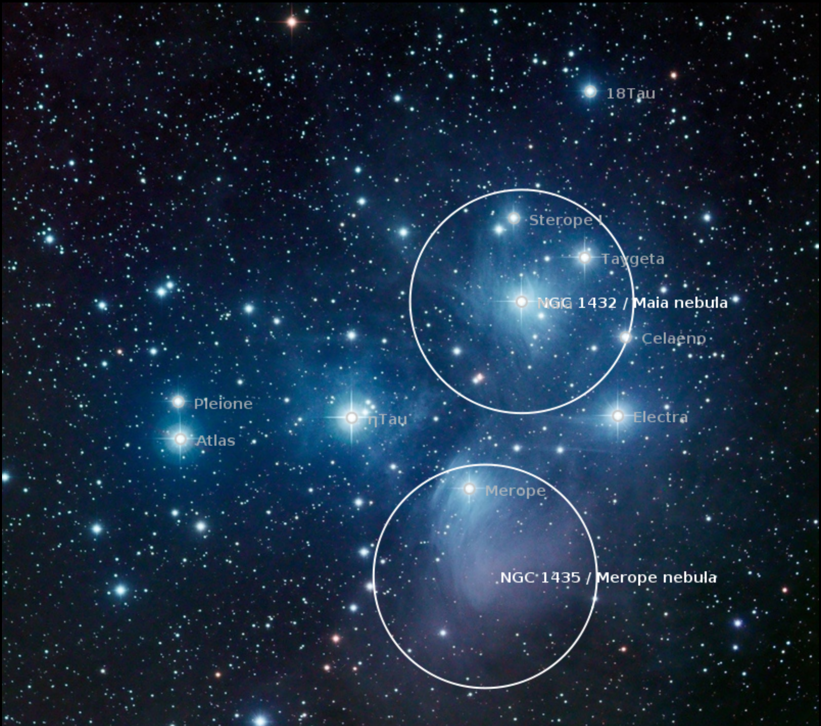
After we find a correct match, we can run a "Tweak" process which tries to fine-tune the World Coordinate System transformation by computing polynomial SIP correction terms. Sometimes this process actually makes the solution worse - we are working on an improved version.

☐ Tweak

Tweak polynomial order: 2

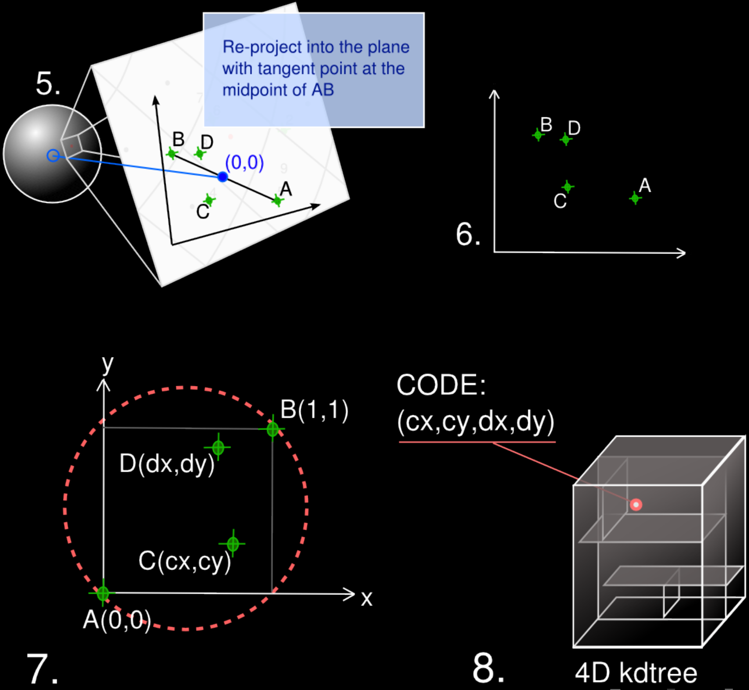


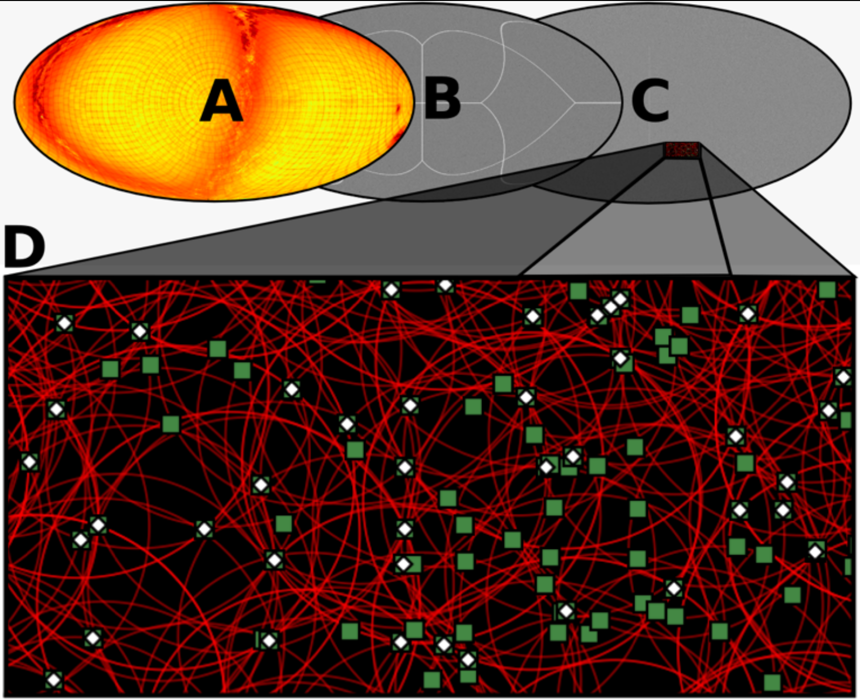



18Tau
Sterope I
Taygeta
NGC 1432 / Maia nebula
Celaeno
Pleione
ηTau
Electra
Atlas
Merope
NGC 1435 / Merope nebula

blind astrometric calibration

- determine astrometric calibration from image pixels *alone*
- geometric hashing step, verification step
- fastest kd-tree in the world (some restrictions apply)
- returns truly *trustworthy* meta-data


5.
Re-project into the plane
with tangent point at the
midpoint of AB
B
D
(0,0)
A
C
6.
B
D
C
A
7.
y
x
B(1,1)
D(dx,dy)
C(cx,cy)
A(0,0)
CODE:
(cx,cy,dx,dy)
8.
4D kdtree


A
B
C
D

performance

- essentially no false positives (exceptions insane)
- $>$ 99.9 percent success rate on *SDSS* and *GALEX* imaging
- large numbers of users (amateurs, professionals, educators, robots)
- can also do other kinds of calibration:
  - wavelength bandpass
  - photometric sensitivity
  - point-spread function
  - date (to within years 2008 AJ 136 1490)
- don't believe? Sign up now at http://astrometry.net/

trusting data

- data that have not been used to *do science* are *wrong data*
- a practical point, not a theoretical point (see *SDSS*)
- the "Virtual Observatory" has no trust model
- science is the *ultimate functional testing environment*

maximally heterogeneous data

- we would like to do simultaneous work with data from amateurs and archives
- data often have unknown provenance, wrong clocks, *etc.*
- we need to calibrate, vet, verify—automatically
- we need *robots that do science*

the theory of everything

- simultaneous modeling of all astronomical imaging (arXiv/0810.3851)
- run in real time as an update system
- model parameters:
  - position and brightness of every star
  - pointing, orientation, bandpass, PSF, calibration of every image
  - camera parameters for every telescope + camera
- report "novel information content" about incoming imaging
- basis of the Open-Source Sky Survey
- calibration and vetting is produced naturally

summary

- doing automated calibration successfully *now*
- this creates a possible *trust* system for astronomy
- we would like to be doing *automated science*
- we have a model for how that might work